GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 13, 2001, 13:35:35 ; Search time 20.6 Seconds
(without alignments)
1550.915 Million cell updates/sec

Title: US-09-784-340-2
Perfect score: 527
Sequence: 1 MRSDKSALVFLLLQLFCVGC.....KCFLFSCQKFNKTRKIEKRE 527

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_0601:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	174	33.0	529	21	AAB28677	Human carbohydrate
2	78	14.8	78	21	AAG03280	Human secreted pro
3	33	6.3	528	21	AAY78933	Human UDP-glucuron
4	33	6.3	530	19	AAW47126	Uridine diphospho-
5	28	5.3	524	21	AAY78934	Human UDP-glucuron
6	23	4.4	530	21	AAY78935	Human UDP-glucuron
7	9	1.7	44	13	AAR30148	UGT1 Exon 2 produc
8	9	1.7	94	21	AAB53721	Human colon cancer
9	9	1.7	98	13	AAR30166	UGT1 Exon 5 produc
10	9	1.7	129	20	AAY29525	Human lung tumour
11	9	1.7	129	21	AAB44411	Human lung tumour-
12	9	1.7	245	21	AAY57100	UDP-glucuronosyltr
13	9	1.7	533	13	AAR26153	HUG-Br1. Homo sap
14	9	1.7	534	13	AAR26154	HUG-Br2. Homo sap
15	8	1.5	68	21	AAB56504	Human prostate can
16	8	1.5	253	21	AAY57099	UDP-glucuronosyltr
17	8	1.5	310	21	AAY57098	UDP-glucuronosyltr
18	8	1.5	317	21	AAY57097	UDP-glucuronosyltr
19	8	1.5	380	14	AAR44512	Elk PTK. Rattus r
20	8	1.5	466	18	AAW09825	UDP-glucose:thiohy
21	8	1.5	951	16	AAR75704	Eph-related CEK6.
22	8	1.5	984	14	AAR44513	elk. Rattus rattu
23	7	1.3	17	13	AAR26152	Transferase conser
24	7	1.3	30	13	AAR25213	Immunosuppressive
25	7	1.3	81	20	AAY07220	Presenilin/beta-am
26	7	1.3	86	21	AAG01230	Human secreted pro
27	7	1.3	88	22	AAB76802	Corynebacterium gl
28	7	1.3	89	21	AAB38042	Fragment of human
29	7	1.3	99	22	AAB11484	Human presenilin p
30	7	1.3	140	20	AAY60184	Human endometrium
31	7	1.3	176	21	AAB32877	Pinus radiata tran
32	7	1.3	176	21	AAB33265	Pinus radiata tran
33	7	1.3	180	21	AAY68872	Amino acid sequenc
34	7	1.3	211	21	AAB37993	Human secreted pro
35	7	1.3	235	20	AAY19973	B. burgdorferi ant
36	7	1.3	244	20	AAY36910	Protein involved i
37	7	1.3	261	20	AAY19972	B. burgdorferi ant
38	7	1.3	303	20	AAY19903	B. burgdorferi ant
39	7	1.3	309	20	AAY37181	Protein involved i
40	7	1.3	322	20	AAY19902	B. burgdorferi ant
41	7	1.3	348	14	AAR41346	Human CAR receptor
42	7	1.3	348	18	AAW32536	Constitutively act
43	7	1.3	348	20	AAW93902	Human CAR receptor
44	7	1.3	357	20	AAY17872	Mouse nuclear rece
45	7	1.3	358	20	AAW93903	Mouse CAR receptor

ALIGNMENTS

RESULT 1
AAB28677
ID AAB28677 standard; Protein; 529 AA.
XX
AC AAB28677;
XX
DT 13-FEB-2001 (first entry)
XX
DE Human carbohydrate-modifying enzyme Incyte ID No: 2912330CD1.
XX
KW Human; carbohydrate-modifying enzyme; CME; antidiabetic;
KW immunosuppressive; anti-HIV; antiinflammatory; antianaemic;
KW antiasthmatic; antiarteriosclerotic; antithyroid; hepatotropic;
KW nephrotropic; antigout; thyromimetic; neuroprotective; osteopathic;
KW antiarthritic; antipsoriatic; uropathic; ophthalmological;
KW dermatological; antiulcer; cytostatic; virucide; antibacterial;
KW fungicide; protozoacide; tranquiliser; vulnerary; diabetes;
KW autoimmune disorder; inflammatory disorder; infection.
XX
OS Homo sapiens.
XX
PN WO200063351-A2.
XX
PD 26-OCT-2000.
XX
PF 20-APR-2000; 2000WO-US10882.
XX
PR 21-APR-1999; 99US-0130383.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lal P, Yue H, Tang YT, Hillman JL, Baughn MR, Yang J;
XX
DR WPI; 2000-672729/65.

```
DR   N-PSDB; AAC65396.
XX
PT   Novel carbohydrate modifying enzyme polypeptides and polynucleotides
PT   for diagnosis, treatment, and prevention of carbohydrate metabolism
PT   disorders, autoimmune/inflammatory disorders, and cancer -
XX
PS   Claim 1; Page 71-72; 75pp; English.
XX
CC   The present sequence is a human carbohydrate-modifying enzyme
CC   (CME). CME polynucleotides and polypeptides are useful for treating and
CC   diagnosing diseases associated with CME such as diabetes,
CC   autoimmune/inflammatory disorders such as AIDS, Addison's disease,
CC   adult respiratory distress syndrome, allergies, anaemia, asthma,
CC   atherosclerosis, autoimmune thyroiditis, bronchitis, cholocystitis,
CC   contact dermatitis, Crohn's disease, emphysema, erythroblastosis fetalis,
CC   glomerulonephritis, Good pasture's syndrome, gout, Grave's disease,
CC   Hashimoto's thyroiditis, multiple sclerosis, myasthenia gravis,
CC   osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis,
CC   Reiter's syndrome, arthritis, scleroderma, sjogren's syndrome, systemic
CC   lupus erythematosus, ulcerative colitis, uveitis, Werner syndrome,
CC   complications of cancer, haemodialysis, and extracorporeal circulation,
CC   viral, bacterial, fungal parasitic, protozoal, and helminthic infections,
CC   trauma, or cancer. CME, or its catalytic or immunogenic fragment, is
CC   useful for drug screening.
XX
SQ   Sequence   529 AA;

  Query Match             33.0%;  Score 174;  DB 21;  Length 529;
  Best Local Similarity   100.0%;  Pred. No. 7.1e-163;
  Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     354 WIPQNDLLGHPKTKAFITHGGMNGIYEAIYHGVPMVGVPIFGDQLDNIAHMKAKGAAVEI 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     356 wipqndllghpktkafithggmngiyeaiyhgvpmvgvpifgdqldniahmkakgaavei 415

Qy     414 NFKTMTSEDLLRALRTVITDSSYKENAMRLSRIHHDQPVKPLDRAVFWIEFVMRHKGAKH 473
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     416 nfktmtsedllralrtvitdssykenamrlsrihhdqpvkpldravfwiefvmrhkgakh 475

Qy     474 LRSAAHDLTWFQHYSIDVIGFLLTCVATAIFLFTKCFLFSCQKFNKTRKIEKRE 527
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     476 lrsaahdltwfqhysidvigflltcvataiflftkcflfscqkfnktrkiekre 529


RESULT   2
AAG03280
ID   AAG03280 standard; Protein; 78 AA.
XX
AC   AAG03280;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 7361.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC03286.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 13; SEQ ID 7361; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
CC   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC   were prepared from total human RNAs or polyA+ RNAs derived from 30
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   78 AA;

  Query Match             14.8%;  Score 78;  DB 21;  Length 78;
  Best Local Similarity   100.0%;  Pred. No. 4.7e-69;
  Matches   78;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     290 MENFVQSSGEDGIVVFSLGSLFQNVTEEKANIIASALAQIPQKVLWRYKGKKPSTLGANT 349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 menfvqssgedgivvfslgslfqnvteekaniiasalaqipqkvlwrykgkkpstlgant 60

Qy     350 RLYDWIPQNDLLGHPKTK 367
           ||||||||||||||||||
Db      61 rlydwipqndllghpktk 78


RESULT   3
AAY78933
ID   AAY78933 standard; Protein; 528 AA.
XX
AC   AAY78933;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Human UDP-glucuronosyltransferase 2B4 amino acid sequence.
XX
KW   UDP-glucuronosyltransferase 2B4; UGT2B4; polymorphism; metabolism; SNPs;
KW   drug interaction; detect; human; single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
PN   WO200006776-A1.
XX
PD   10-FEB-2000.
XX
PF   22-JUL-1999;   99WO-US16675.
XX
PR   28-JUL-1998;   98US-0094391.
XX
PA   (AXYS-) AXYS PHARM INC.
XX
PI   Galvin M,  Miller A,  Penny L,  Riedy M;
XX
DR   WPI; 2000-195321/17.
DR   N-PSDB; AAZ95199.
XX
PT   Novel human UDP-glucuronosyltransferase sequence, polymorphisms for
PT   genotyping individuals to predict rate of metabolism of substrates and
PT   for identifying potential drug interactions -
XX
PS   Disclosure; Page 36-37; 72pp; English.
```

```
XX
CC   This sequence represents the human UDP-glucuronosyltransferase 2B4
CC   (UGT2B4) amino acid sequence. UDP-glucuronosyltransferase (UGTs) are a
CC   family of enzymes that catalyse the glucuronic acid conjugation of a
CC   wide range of endogenous and exogenous substrates. The UGT2B gene
CC   subfamily encode steroid metabolizing isoforms in the liver. Alteration
CC   of the expression or function of UGTs may effect drug metabolism. The
CC   invention relates to non-chromosomal nucleic acid molecules, which
CC   comprise human UGT2B sequence polymorphisms (see AAZ95051-Z95110). Probes
CC   which detect the UGT2B locus polymorphisms can be used to detect altered
CC   UGT2B metabolism of a substrate in an individual. The nucleic acid
CC   molecules comprising a human UGT2B sequence polymorphism can be used in
CC   screening assays for genotyping individuals, also to predict their rate
CC   of metabolism of UGT2B substrate, potential drug-drug interactions and
CC   adverse side effects. The polymorphisms can be used as single nucleotide
CC   polymorphisms (SNPs) for detecting genetic linkage related to phenotypic
CC   variation in activity or expression of UGT2B protein. The polymorphism
CC   containing nucleic acid molecules may also be used for generating
CC   genetically modified non-human animals and for obtaining site specific
CC   gene modification in cell lines.
XX
SQ   Sequence   528 AA;


  Query Match             6.3%;  Score 33;  DB 21;  Length 528;
  Best Local Similarity   100.0%;  Pred. No. 6e-24;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     443 LSRIHHDQPVKPLDRAVFWIEFVMRHKGAKHLR 475
           |||||||||||||||||||||||||||||||||
Db     445 lsrihhdqpvkpldravfwiefvmrhkgakhlr 477


RESULT   4
AAW47126
ID   AAW47126 standard; Protein; 530 AA.
XX
AC   AAW47126;
XX
DT   26-MAY-1998  (first entry)
XX
DE   Uridine diphospho-glucuronosyltransferase 2B17 (UGT2B17) enzyme.
XX
KW   Uridine diphospho-glucuronosyltransferase 2B17; UGT2B17; catalyse;
KW   androsterone; androsterone-glucuronic acid; androgen; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO9744466-A1.
XX
PD   27-NOV-1997.
XX
PF   16-MAY-1997;   97WO-CA00328.
XX
PR   17-MAY-1996;   96US-0649319.
XX
PA   (ENDO-) ENDORECHERCHE INC.
XX
PI   Beaulieu M,  Belanger A,  Hum DW,  Levesque E;
XX
DR   WPI; 1998-018520/02.
DR   N-PSDB; AAV15900.
XX
PT   DNA encoding uridine di:phospho:glucuronosyl:transferase 2B17 -
PT   which catalyses conversion of androsterone to
PT   androsterone-glucuronic acid
XX
PS   Claim 16; Pages 4-6; 53pp; English.
XX
CC   This is the enzyme uridine diphospho-glucuronosyltransferase 2B17
CC   (UGT2B17). This novel enzyme catalyses the conversion of androsterone
CC   to androsterone-glucuronic acid. The UGT2B17 can be used to detect
CC   anti-UGT2B17 antibodies. The antibody can be used to detect a localised
CC   concentration of UGT2B17 or an alteration in androgen activity. The
CC   UGT2B17 can also be used to alter the concentration of an androgenic
CC   compound in a tissue, specifically dihydrotestosterone. An isolated
CC   nucleotide sequence comprising at least 30 consecutive nucleotides from
CC   the coding region of the 2107 base pair sequence, or its complement can
CC   be used to block the synthesis of UGT2B17, e.g. an expression disrupting
CC   sense or antisense fragment, or as a probe for a UGT2B17 coding sequence.
XX
SQ   Sequence   530 AA;


  Query Match             6.3%;  Score 33;  DB 19;  Length 530;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-24;
  Matches   33;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     443 LSRIHHDQPVKPLDRAVFWIEFVMRHKGAKHLR 475
           |||||||||||||||||||||||||||||||||
Db     446 lsrihhdqpvkpldravfwiefvmrhkgakhlr 478


RESULT   5
AAY78934
ID   AAY78934 standard; Protein; 524 AA.
XX
AC   AAY78934;
XX
DT   05-JUN-2000  (first entry)
XX
DE   Human UDP-glucuronosyltransferase 2B7 amino acid sequence.
XX
KW   UDP-glucuronosyltransferase 2B4; UGT2B4; polymorphism; metabolism; SNPs;
KW   drug interaction; detect; human; single nucleotide polymorphism.
XX
OS   Homo sapiens.
XX
PN   WO200006776-A1.
XX
PD   10-FEB-2000.
XX
PF   22-JUL-1999;   99WO-US16675.
XX
PR   28-JUL-1998;   98US-0094391.
XX
PA   (AXYS-) AXYS PHARM INC.
XX
PI   Galvin M,  Miller A,  Penny L,  Riedy M;
XX
DR   WPI; 2000-195321/17.
DR   N-PSDB; AAZ95200.
XX
PT   Novel human UDP-glucuronosyltransferase sequence, polymorphisms for
PT   genotyping individuals to predict rate of metabolism of substrates and
PT   for identifying potential drug interactions -
XX
PS   Disclosure; Page 44-45; 72pp; English.
XX
CC   This sequence represents the human UDP-glucuronosyltransferase 2B7
CC   (UGT2B7) amino acid sequence. UDP-glucuronosyltransferase (UGTs) are a
CC   family of enzymes that catalyse the glucuronic acid conjugation of a
CC   wide range of endogenous and exogenous substrates. The UGT2B gene
CC   subfamily encode steroid metabolizing isoforms in the liver. Alteration
CC   of the expression or function of UGTs may effect drug metabolism. The
CC   invention relates to non-chromosomal nucleic acid molecules, which
CC   comprise human UGT2B sequence polymorphisms (see AAZ95051-Z95110). Probes
CC   which detect the UGT2B locus polymorphisms can be used to detect altered
CC   UGT2B metabolism of a substrate in an individual. The nucleic acid
CC   molecules comprising a human UGT2B sequence polymorphism can be used in
CC   screening assays for genotyping individuals, also to predict their rate
CC   of metabolism of UGT2B substrate, potential drug-drug interactions and
CC   adverse side effects. The polymorphisms can be used as single nucleotide
CC   polymorphisms (SNPs) for detecting genetic linkage related to phenotypic
CC   variation in activity or expression of UGT2B protein. The polymorphism
CC   containing nucleic acid molecules may also be used for generating
```

CC genetically modified non-human animals and for obtaining site specific
CC gene modification in cell lines.
XX
SQ Sequence 524 AA;

Query Match 5.3%; Score 28; DB 21; Length 524;
Best Local Similarity 100.0%; Pred. No. 5.1e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 448 HDQPVKPLDRAVFWIEFVMRHKGAKHLR 475
||||||||||||||||||||||||||||
Db 450 hdqpvkpldravfwiefvmrhkgakhlr 477

RESULT 6
AAY78935
ID AAY78935 standard; Protein; 530 AA.
XX
AC AAY78935;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human UDP-glucuronosyltransferase 2B15 amino acid sequence.
XX
KW UDP-glucuronosyltransferase 2B15; UGT2B15; polymorphism; metabolism;
KW drug interaction; detect; human; single nucleotide polymorphism; SNPs.
XX
OS Homo sapiens.
XX
PN WO200006776-A1.
XX
PD 10-FEB-2000.
XX
PF 22-JUL-1999; 99WO-US16675.
XX
PR 28-JUL-1998; 98US-0094391.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Galvin M, Miller A, Penny L, Riedy M;
XX
DR WPI; 2000-195321/17.
DR N-PSDB; AAZ95206.
XX
PT Novel human UDP-glucuronosyltransferase sequence, polymorphisms for
PT genotyping individuals to predict rate of metabolism of substrates and
PT for identifying potential drug interactions -
XX
PS Disclosure; Page 59-60; 72pp; English.
XX
CC This sequence represents the human UDP-glucuronosyltransferase 2B15
CC (UGT2B15) amino acid sequence. UDP-glucuronosyltransferase (UGTs) are a
CC family of enzymes that catalyse the glucuronic acid conjugation of a
CC wide range of endogenous and exogenous substrates. The UGT2B gene
CC subfamily encode steroid metabolizing isoforms in the liver. Alteration
CC of the expression or function of UGTs may effect drug metabolism. The
CC invention relates to non-chromosomal nucleic acid molecules, which
CC comprise human UGT2B sequence polymorphisms (see AAZ95051-Z95110). Probes
CC which detect the UGT2B locus polymorphisms can be used to detect altered
CC UGT2B metabolism of a substrate in an individual. The nucleic acid
CC molecules comprising a human UGT2B sequence polymorphism can be used in
CC screening assays for genotyping individuals, also to predict their rate
CC of metabolism of UGT2B substrate, potential drug-drug interactions and
CC adverse side effects. The polymorphisms can be used as single nucleotide
CC polymorphisms (SNPs) for detecting genetic linkage related to phenotypic
CC variation in activity or expression of UGT2B protein. The polymorphism
CC containing nucleic acid molecules may also be used for generating
CC genetically modified non-human animals and for obtaining site specific
CC gene modification in cell lines.
XX
SQ Sequence 530 AA;

Query Match 4.4%; Score 23; DB 21; Length 530;
Best Local Similarity 100.0%; Pred. No. 4.3e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 453 KPLDRAVFWIEFVMRHKGAKHLR 475
|||||||||||||||||||||||
Db 456 kpldravfwiefvmrhkgakhlr 478

RESULT 7
AAR30148
ID AAR30148 standard; Protein; 44 AA.
XX
AC AAR30148;
XX
DT 27-JAN-1993 (first entry)
XX
DE UGT1 Exon 2 product.
XX
KW UGT1A; UGT1BP; UGT1C; UGT1D; UGT1E; UGT1F; isozyme; bilirubin;
KW UDP-glucuronosyl transferase; CN.
XX
OS Homo sapiens.
XX
PN WO9212987-A.
XX
PD 06-AUG-1992.
XX
PF 10-JAN-1992; 92WO-US00282.
XX
PR 10-JAN-1991; 91US-0639453.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Owens IS, Ritter JK;
XX
DR WPI; 1992-284593/34.
DR N-PSDB; AAQ33025.
XX
PT Isolated gene locus UGT1, DNA segments and diagnostic probes -
PT for diagnosing Gilbert's disease and Crigler-Najjar syndrome
PT types I and II
XX
PS Disclosure; Fig 1G; 99pp; English.
XX
CC The isolated gene locus, UGT1, has a sequence of about 10000 bp
CC which represent (1) Exon 1, comprising 6 transcriptional units
CC (UGT1F, E, D, C, BP and A), represented in AAQ27368 and
CC AAQ33020-24 respectively;
CC (2) Exon 2, represented in AAQ33025;
CC (3) Exon 3, represented in AAQ33026;
CC (4) Exon 4, represented in AAQ33026;
CC (5) Exon 5, represented in AAQ33027; and
CC (6) about 69 kb of non-sequenced DNA.
CC Six unique N-termini of 286-289 amino acids are encoded by
CC the six different first exons and identical C-termini of 246 amino
CC acids are encoded by the common exons 2-5. The UGT1 gene locus
CC encodes a family of UDP-glucuronosyl transferase isozymes, two of
CC which metabolise bilirubin.
CC Patients having Crigler-Najjar Syndrome (CN) Type I, have a
CC mutation present in the second common exon.
XX
SQ Sequence 44 AA;

Query Match 1.7%; Score 9; DB 13; Length 44;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 301 GIVVFSLGS 309
|||||||||
Db 13 givvfslgs 21

```
RESULT   8
AAB53721
ID   AAB53721 standard; Protein; 94 AA.
XX
AC   AAB53721;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human colon cancer antigen protein sequence SEQ ID NO:1261.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW   immunomodulatory; muscular; gynaecological; gastrointestinal;
KW   nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; renal disorder;
KW   infectious disease; cardiovascular disorder.
XX
OS   Homo sapiens.
XX
PN   WO200055351-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05883.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587534/55.
DR   N-PSDB; AAC98478.
XX
PT   Colon cancer associated gene sequences, referred to as colon cancer
PT   antigens, useful for the treatment, prevention, and diagnosis of colon
PT   disorders such as colon cancer -
XX
PS   Claim 11; Page 1849; 2104pp; English.
XX
CC   AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC   called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC   human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC   neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC   vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC   can be used in gene therapy. The colon cancer antigen polynucleotides,
CC   proteins and antibodies to the proteins are useful for the prevention,
CC   treatment and diagnosis of colon disorders, such as colon cancer. The
CC   polynucleotides may be used in diagnostics and research, such as for
CC   chromosome identification, and as hybridisation probes. The proteins
CC   may also be used to prevent diseases such as neural disorders, immune
CC   system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, wounds, renal disorders, infectious
CC   diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC   AAB54007 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   94 AA;


  Query Match             1.7%;  Score 9;  DB 21;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 0.55;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    463 EFVMRHKGA 471
          |||||||||
Db     66 efvmrhkga 74


RESULT   9
AAR30166
ID   AAR30166 standard; Protein; 98 AA.
XX
AC   AAR30166;
XX
DT   27-JAN-1993  (first entry)
XX
DE   UGT1 Exon 5 product.
XX
KW   UGT1A; UGT1BP; UGT1C; UGT1D; UGT1E; UGT1F; isozyme; bilirubin;
KW   UDP-glucuronosyl transferase; CN.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 64
FT                   /note= "Val encoded by TGG!"
XX
PN   WO9212987-A.
XX
PD   06-AUG-1992.
XX
PF   10-JAN-1992;   92WO-US00282.
XX
PR   10-JAN-1991;   91US-0639453.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI   Owens IS,  Ritter JK;
XX
DR   WPI; 1992-284593/34.
DR   N-PSDB; AAQ33027.
XX
PT   Isolated gene locus UGT1, DNA segments and diagnostic probes -
PT   for diagnosing Gilbert's disease and Crigler-Najjar syndrome
PT   types I and II
XX
PS   Disclosure; Fig 1I; 99pp; English.
XX
CC   In order to obtain this amino acid sequence, base G485 of
CC   the of the encoding sequence of AAQ33027 needed to be deleted.
CC   The isolated gene locus, UGT1, has a sequence of about 10000 bp
CC   which represent (1) Exon 1, comprising 6 transcriptional units
CC   (UGT1F, E, D, C, BP and A), represented in AAQ27368 and
CC   AAQ33020-24 respectively;
CC                   (2) Exon 2, represented in AAQ33025;
CC                   (3) Exon 3, represented in AAQ33026;
CC                   (4) Exon 4, represented in AAQ33026;
CC                   (5) Exon 5, represented in AAQ33027; and
CC                   (6) about 69 kb of non-sequenced DNA.
CC   Six unique N-termini of 286-289 amino acids are encoded by
CC   the six different first exons and identical C-termini of 246 amino
CC   acids are encoded by the common exons 2-5.  The UGT1 gene locus
CC   encodes a family of UDP-glucuronosyl transferase isozymes, two of
CC   which metabolise bilirubin.
CC   Patients having Crigler-Najjar Syndrome (CN) Type I, have a
CC   mutation present in the second common exon.
XX
SQ   Sequence   98 AA;


  Query Match             1.7%;  Score 9;  DB 13;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 0.57;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    463 EFVMRHKGA 471
          |||||||||
Db     28 efvmrhkga 36


RESULT  10
AAY29525
ID   AAY29525 standard; Protein; 129 AA.
```

XX
AC AAY29525;
XX
DT 13-OCT-1999 (first entry)
XX
DE Human lung tumour protein LT86-5 predicted amino acid sequence.
XX
KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
KW immunotherapy; detection; inhibition.
XX
OS Homo sapiens.
XX
PN WO9938973-A2.
XX
PD 05-AUG-1999.
XX
PF 26-JAN-1999; 99WO-US01642.
XX
PR 22-DEC-1998; 98US-0219245.
PR 28-JAN-1998; 98US-0015022.
PR 28-JAN-1998; 98US-0015029.
PR 18-MAR-1998; 98US-0040828.
PR 18-MAR-1998; 98US-0040831.
PR 23-JUL-1998; 98US-0122191.
PR 23-JUL-1998; 98US-0122192.
XX
PA (CORI-) CORIXA CORP.
XX
PI Frudakis TN, Lodes MJ, Mohamath R, Reed SG;
XX
DR WPI; 1999-479187/40.
DR N-PSDB; AAZ07208.
XX
PT Lung tumour specific polynucleotides for inhibiting the development
PT of lung cancer
XX
PS Example 2; Page 73; 171pp; English.
XX
CC The present invention describes lung tumour specific polynucleotides
CC and tumour antigens. AAZ07144 to AAZ07246 and AAZ08301 to AAZ08325
CC represent specifically claimed polynucleotides, and AAY29486 to AAY29571
CC represent amino acid sequences from the present invention. The lung
CC tumour specific polynucleotides and polypeptides can be used in
CC pharmaceutical compositions and vaccines to inhibit the development of
CC lung cancer. They can also be used to detect lung cancer in a patient.
CC Probes and antibodies derived from the lung tumour sequences are useful
CC in detection of lung cancer.
XX
SQ Sequence 129 AA;

Query Match 1.7%; Score 9; DB 20; Length 129;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 301 GIVVFSLGS 309
|||||||||
Db 1 givvfslgs 9

RESULT 11
AAB44411
ID AAB44411 standard; Protein; 129 AA.
XX
AC AAB44411;
XX
DT 05-FEB-2001 (first entry)
XX
DE Human lung tumour-specific antigen encoded by cDNA #21.
XX
KW Lung tumour protein; lung cancer; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN WO200060077-A2.
XX
PD 12-OCT-2000.
XX
PF 30-MAR-2000; 2000WO-US08560.
XX
PR 02-APR-1999; 99US-0285323.
PR 09-AUG-1999; 99US-0370838.
PR 30-DEC-1999; 99US-0476235.
PR 03-MAR-2000; 2000US-0518809.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H;
XX
DR WPI; 2000-638466/61.
DR N-PSDB; AAC79066.
XX
PT Novel lung tumor polypeptides and polynucleotides, useful for
PT detecting, monitoring or treating cancer, especially lung cancer -
XX
PS Claim 1; Page 99; 243pp; English.
XX
CC The present sequence is given in a specification relating to compounds
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC least an immunogenic part of a lung tumour protein are disclosed.
CC The polypeptides are useful for inhibiting the development of cancer,
CC especially lung cancer. Samples of T cells expressing the polypeptides
CC may be used to inhibit the development of cancer. The polypeptides are
CC also useful for detecting and monitoring the progression of cancer,
CC especially lung cancer.
XX
SQ Sequence 129 AA;

Query Match 1.7%; Score 9; DB 21; Length 129;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 301 GIVVFSLGS 309
|||||||||
Db 1 givvfslgs 9

RESULT 12
AAY57100
ID AAY57100 standard; Protein; 245 AA.
XX
AC AAY57100;
XX
DT 28-FEB-2000 (first entry)
XX
DE UDP-glucuronosyltransferase 1 (UGT1) exons 2-5 amino acid sequence.
XX
KW Uridine diphosphate-glucuronosyltransferase 1; UGT1; polymorphism; probe;
KW glucuronic acid; Crigler-Najjar syndrome; Gilbert syndrome; jaundice;
KW unconjugated hyperbilirubinaemia; drug metabolism; transgenic animal;
KW pharmacogenetic screening; diagnose.
XX
OS Homo sapiens.
XX
PN WO9957322-A2.
XX
PD 11-NOV-1999.
XX
PF 04-MAY-1999; 99WO-US09702.
XX
PR 07-MAY-1998; 98US-0084807.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Penny L, Galvin M;

```
XX
DR   WPI; 2000-052981/04.
DR   N-PSDB; AAZ45118.
XX
PT   New nucleic acid representing polymorphisms in the human uridine
PT   diphosphate glucuronosyltransferase gene, used for diagnosis and
PT   evaluation of drug metabolism -
XX
PS   Examples; Page 44-45; 63pp; English.
XX
CC   AAY57092-Y57100 are the amino acid sequences of exons 1A-1J of human
CC   uridine diphosphate-glucuronosyltransferase 1 (UGT1). The UGTs are a
CC   family of enzymes that catalyse the glucuronic acid conjugation of a
CC   wide range of endogenous and exogenous substrates including phenols,
CC   alcohols, amines and fatty acids. Many of the reactions catalysed by
CC   UGTs result in toxic substances being converted to compounds which are
CC   more water soluble and are excreted. The invention relates to and
CC   identifies UGT1 polymorphisms (AAZ45004-Z45041). The polymorphism
CC   sequences are useful as probes for detecting UGT1 locus polymorphisms,
CC   indicative of altered UGT1 expression or activity. These polymorphisms
CC   are associated with Crigler-Najjar and Gilbert syndromes (unconjugated
CC   hyperbilirubinaemia) and drug metabolism. The genotyping of the UGT1 gene
CC   is used to predict the rate of metabolism of UGT1 substrates, possible
CC   drug-drug interactions and adverse side effects (i.e. to optimize drug
CC   dosage), and to screen for diseases caused by exposure to toxins and to
CC   study the effects of polymorphisms on enzymatic activity. The UGT1
CC   sequences, including polymorphisms, can also be used to produce the
CC   corresponding protein (or its fragments) or to generate transgenic
CC   animals or modified cells e.g. for pharmacogenetic screening.
XX
SQ   Sequence   245 AA;


  Query Match             1.7%;  Score 9;  DB 21;  Length 245;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     301 GIVVFSLGS 309
           |||||||||
Db      13 givvfslgs 21


RESULT  13
AAR26153
ID   AAR26153 standard; Protein; 533 AA.
XX
AC   AAR26153;
XX
DT   27-JAN-1993  (first entry)
XX
DE   HUG-Br1.
XX
KW   Bilirubin; UDP-glucuronosyltransferase; HUGBr1; HUGBr2;
KW   monoglucuronide; diglucuronide.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          10..20
FT                   /note= "putative membrane-insertion signal"
FT   Region          491..507
FT                   /note= "putative membrane-anchoring peptide"
FT   Modified-site   102
FT                   /note= "predicted Asn-linked glycosylation site"
FT   Modified-site   295
FT                   /note= "predicted Asn-linked glycosylation site"
FT   Modified-site   347
FT                   /note= "predicted Asn-linked glycosylation site"
FT   Misc-difference 158
FT                   /note= "feature not labelled in specification"
FT   Misc-difference 181
FT                   /note= "feature not labelled in specification"
FT   Misc-difference 228
FT                   /note= "feature not labelled in specification"
XX
PN   WO9212987-A.
XX
PD   06-AUG-1992.
XX
PF   10-JAN-1992;   92WO-US00282.
XX
PR   10-JAN-1991;   91US-0639453.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI   Owens IS,  Ritter JK;
XX
DR   WPI; 1992-284593/34.
DR   N-PSDB; AAQ27369.
XX
PT   Isolated gene locus UGT1, DNA segments and diagnostic probes -
PT   for diagnosing Gilbert's disease and Crigler-Najjar syndrome
PT   types I and II
XX
PS   Disclosure; Fig 9A-I; 99pp; English.
XX
CC   Two human liver bilirubin UDP-glucuronosyltransferase cDNAs have
CC   been isolated.  They are referred to as HUGBr1 (AAQ27369) and HUGBr2
CC   (AAQ27370) (Ritter, et al., J. Biol. Chem. 266:1043-1047 (1991)) and,
CC   upon expression individually in COS-1 cells, encode isoforms that
CC   catalyse the formation of the two bilirubin monoglucuronides and
CC   the diglucuronide.
CC   The cDNAs contain identical 3' ends (1469 bp in length) to each
CC   other and to that of the human phenol transferase cDNA, HLUGP1
CC   (Harding et al., Proc. Natl. Aca. Sci. USA 85:8281 (1988)).
CC   In contrast, they have unique 5' ends.
XX
SQ   Sequence   533 AA;


  Query Match             1.7%;  Score 9;  DB 13;  Length 533;
  Best Local Similarity   100.0%;  Pred. No. 2.7;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     356 PQNDLLGHP 364
           |||||||||
Db     356 pqndllghp 364


RESULT  14
AAR26154
ID   AAR26154 standard; Protein; 534 AA.
XX
AC   AAR26154;
XX
DT   27-JAN-1993  (first entry)
XX
DE   HUG-Br2.
XX
KW   Bilirubin; UDP-glucuronosyltransferase; HUGBr1; HUGBr2;
KW   monoglucuronide; diglucuronide.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          12..22
FT                   /note= "putative membrane-insertion signal"
FT   Region          492..508
FT                   /note= "putative membrane-anchoring peptide"
FT   Modified-site   348
FT                   /note= "predicted Asn-linked glycosylation site"
FT   Misc-difference 282..285
FT                   /note= "residues encoded by TGCCAACGGGAAG !"
XX
PN   WO9212987-A.
XX
```

PD 06-AUG-1992.
XX
PF 10-JAN-1992; 92WO-US00282.
XX
PR 10-JAN-1991; 91US-0639453.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Owens IS, Ritter JK;
XX
DR WPI; 1992-284593/34.
DR N-PSDB; AAQ27369.
XX
PT Isolated gene locus UGT1, DNA segments and diagnostic probes -
PT for diagnosing Gilbert's disease and Crigler-Najjar syndrome
PT types I and II
XX
PS Disclosure; Fig 9A-I; 99pp; English.
XX
CC Two human liver bilirubin UDP-glucuronosyltransferase cDNAs have
CC been isolated. They are referred to as HUGBr1 (AAQ27369) and HUGBr2
CC (AAQ27370) (Ritter, et al., J. Biol. Chem. 266:1043-1047 (1991)) and,
CC upon expression individually in COS-1 cells, encode isoforms that
CC catalyse the formation of the two bilirubin monoglucuronides and
CC the diglucuronide.
CC The cDNAs contain identical 3' ends (1469 bp in length) to each
CC other and to that of the human phenol transferase cDNA, HLUGP1
CC (Harding et al., Proc. Natl. Aca. Sci. USA 85:8281 (1988)).
CC In contrast, they have unique 5' ends.
XX
SQ Sequence 534 AA;

Query Match 1.7%; Score 9; DB 13; Length 534;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 356 PQNDLLGHP 364
|||||||||
Db 357 pqndllghp 365

RESULT 15
AAB56504
ID AAB56504 standard; Protein; 68 AA.
XX
AC AAB56504;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen protein sequence SEQ ID NO:1082.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX
OS Homo sapiens.
XX
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587513/55.
DR N-PSDB; AAF15707.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
PS Claim 11; Page 1507; 2338pp; English.
XX
CC AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 68 AA;

Query Match 1.5%; Score 8; DB 21; Length 68;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 CGKVLVWP 30
||||||||
Db 24 cgkvlvwp 31

RESULT 16
AAY57099
ID AAY57099 standard; Protein; 253 AA.
XX
AC AAY57099;
XX
DT 28-FEB-2000 (first entry)
XX
DE UDP-glucuronosyltransferase 1 (UGT1) exon 1J amino acid sequence.
XX
KW Uridine diphosphate-glucuronosyltransferase 1; UGT1; polymorphism; probe;
KW glucuronic acid; Crigler-Najjar syndrome; Gilbert syndrome; jaundice;
KW unconjugated hyperbilirubinaemia; drug metabolism; transgenic animal;
KW pharmacogenetic screening; diagnose.
XX
OS Homo sapiens.
XX
PN WO9957322-A2.
XX
PD 11-NOV-1999.
XX
PF 04-MAY-1999; 99WO-US09702.
XX
PR 07-MAY-1998; 98US-0084807.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Penny L, Galvin M;
XX
DR WPI; 2000-052981/04.
DR N-PSDB; AAZ45117.
XX
PT New nucleic acid representing polymorphisms in the human uridine
PT diphosphate glucuronosyltransferase gene, used for diagnosis and
PT evaluation of drug metabolism -
XX

PS Examples; Page 43; 63pp; English.
XX
CC AAY57092-Y57100 are the amino acid sequences of exons 1A-1J of human
CC uridine diphosphate-glucuronosyltransferase 1 (UGT1). The UGTs are a
CC family of enzymes that catalyse the glucuronic acid conjugation of a
CC wide range of endogenous and exogenous substrates including phenols,
CC alcohols, amines and fatty acids. Many of the reactions catalysed by
CC UGTs result in toxic substances being converted to compounds which are
CC more water soluble and are excreted. The invention relates to and
CC identifies UGT1 polymorphisms (AAZ45004-Z45041). The polymorphism
CC sequences are useful as probes for detecting UGT1 locus polymorphisms,
CC indicative of altered UGT1 expression or activity. These polymorphisms
CC are associated with Crigler-Najjar and Gilbert syndromes (unconjugated
CC hyperbilirubinaemia) and drug metabolism. The genotyping of the UGT1 gene
CC is used to predict the rate of metabolism of UGT1 substrates, possible
CC drug-drug interactions and adverse side effects (i.e. to optimize drug
CC dosage), and to screen for diseases caused by exposure to toxins and to
CC study the effects of polymorphisms on enzymatic activity. The UGT1
CC sequences, including polymorphisms, can also be used to produce the
CC corresponding protein (or its fragments) or to generate transgenic
CC animals or modified cells e.g. for pharmacogenetic screening.
XX
SQ Sequence 253 AA;

Query Match 1.5%; Score 8; DB 21; Length 253;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 186 PAPLSYVP 193
||||||||
Db 152 paplsyvp 159

RESULT 17
AAY57098
ID AAY57098 standard; Protein; 310 AA.
XX
AC AAY57098;
XX
DT 28-FEB-2000 (first entry)
XX
DE UDP-glucuronosyltransferase 1 (UGT1) exon 1H amino acid sequence.
XX
KW Uridine diphosphate-glucuronosyltransferase 1; UGT1; polymorphism; probe;
KW glucuronic acid; Crigler-Najjar syndrome; Gilbert syndrome; jaundice;
KW unconjugated hyperbilirubinaemia; drug metabolism; transgenic animal;
KW pharmacogenetic screening; diagnose.
XX
OS Homo sapiens.
XX
PN WO9957322-A2.
XX
PD 11-NOV-1999.
XX
PF 04-MAY-1999; 99WO-US09702.
XX
PR 07-MAY-1998; 98US-0084807.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Penny L, Galvin M;
XX
DR WPI; 2000-052981/04.
DR N-PSDB; AAZ45116.
XX
PT New nucleic acid representing polymorphisms in the human uridine
PT diphosphate glucuronosyltransferase gene, used for diagnosis and
PT evaluation of drug metabolism -
XX
PS Examples; Page 41; 63pp; English.
XX
CC AAY57092-Y57100 are the amino acid sequences of exons 1A-1J of human
CC uridine diphosphate-glucuronosyltransferase 1 (UGT1). The UGTs are a
CC family of enzymes that catalyse the glucuronic acid conjugation of a
CC wide range of endogenous and exogenous substrates including phenols,
CC alcohols, amines and fatty acids. Many of the reactions catalysed by
CC UGTs result in toxic substances being converted to compounds which are
CC more water soluble and are excreted. The invention relates to and
CC identifies UGT1 polymorphisms (AAZ45004-Z45041). The polymorphism
CC sequences are useful as probes for detecting UGT1 locus polymorphisms,
CC indicative of altered UGT1 expression or activity. These polymorphisms
CC are associated with Crigler-Najjar and Gilbert syndromes (unconjugated
CC hyperbilirubinaemia) and drug metabolism. The genotyping of the UGT1 gene
CC is used to predict the rate of metabolism of UGT1 substrates, possible
CC drug-drug interactions and adverse side effects (i.e. to optimize drug
CC dosage), and to screen for diseases caused by exposure to toxins and to
CC study the effects of polymorphisms on enzymatic activity. The UGT1
CC sequences, including polymorphisms, can also be used to produce the
CC corresponding protein (or its fragments) or to generate transgenic
CC animals or modified cells e.g. for pharmacogenetic screening.
XX
SQ Sequence 310 AA;

Query Match 1.5%; Score 8; DB 21; Length 310;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 186 PAPLSYVP 193
||||||||
Db 184 paplsyvp 191

RESULT 18
AAY57097
ID AAY57097 standard; Protein; 317 AA.
XX
AC AAY57097;
XX
DT 28-FEB-2000 (first entry)
XX
DE UDP-glucuronosyltransferase 1 (UGT1) exon 1G amino acid sequence.
XX
KW Uridine diphosphate-glucuronosyltransferase 1; UGT1; polymorphism; probe;
KW glucuronic acid; Crigler-Najjar syndrome; Gilbert syndrome; jaundice;
KW unconjugated hyperbilirubinaemia; drug metabolism; transgenic animal;
KW pharmacogenetic screening; diagnose.
XX
OS Homo sapiens.
XX
PN WO9957322-A2.
XX
PD 11-NOV-1999.
XX
PF 04-MAY-1999; 99WO-US09702.
XX
PR 07-MAY-1998; 98US-0084807.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Penny L, Galvin M;
XX
DR WPI; 2000-052981/04.
DR N-PSDB; AAZ45115.
XX
PT New nucleic acid representing polymorphisms in the human uridine
PT diphosphate glucuronosyltransferase gene, used for diagnosis and
PT evaluation of drug metabolism -
XX
PS Examples; Page 38-39; 63pp; English.
XX
CC AAY57092-Y57100 are the amino acid sequences of exons 1A-1J of human
CC uridine diphosphate-glucuronosyltransferase 1 (UGT1). The UGTs are a
CC family of enzymes that catalyse the glucuronic acid conjugation of a
CC wide range of endogenous and exogenous substrates including phenols,

```
CC   alcohols, amines and fatty acids. Many of the reactions catalysed by
CC   UGTs result in toxic substances being converted to compounds which are
CC   more water soluble and are excreted. The invention relates to and
CC   identifies UGT1 polymorphisms (AAZ45004-Z45041). The polymorphism
CC   sequences are useful as probes for detecting UGT1 locus polymorphisms,
CC   indicative of altered UGT1 expression or activity. These polymorphisms
CC   are associated with Crigler-Najjar and Gilbert syndromes (unconjugated
CC   hyperbilirubinaemia) and drug metabolism. The genotyping of the UGT1 gene
CC   is used to predict the rate of metabolism of UGT1 substrates, possible
CC   drug-drug interactions and adverse side effects (i.e. to optimize drug
CC   dosage), and to screen for diseases caused by exposure to toxins and to
CC   study the effects of polymorphisms on enzymatic activity. The UGT1
CC   sequences, including polymorphisms, can also be used to produce the
CC   corresponding protein (or its fragments) or to generate transgenic
CC   animals or modified cells e.g. for pharmacogenetic screening.
XX
SQ   Sequence   317 AA;


  Query Match             1.5%;  Score 8;  DB 21;  Length 317;
  Best Local Similarity   100.0%;  Pred. No. 16;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     186 PAPLSYVP 193
           ||||||||
Db     184 paplsyvp 191


RESULT  19
AAR44512
ID   AAR44512 standard; Protein; 380 AA.
XX
AC   AAR44512;
XX
DT   16-JUN-1994  (first entry)
XX
DE   Elk PTK.
XX
KW   Lambda gt11; expression vector; lambda-B1-Elk; protein tyrosine kinase;
KW   Elk; B1; Eph; subfamily; receptor-like tyrosine kinase; eph; eck;
KW   phosphorylation; phosphorylated kinase insert domain; growth factor;
KW   receptor kinase; platelet-derived growth factor receptor.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   Region          22
FT                   /note= "PTK conserved amino acid"
FT   Region          24
FT                   /note= "PTK conserved amino acid"
FT   Region          27
FT                   /note= "PTK conserved amino acid"
FT   Region          29
FT                   /note= "PTK conserved amino acid"
FT   Region          45
FT                   /note= "PTK conserved amino acid"
FT   Region          47
FT                   /note= "PTK conserved amino acid"
FT   Region          64
FT                   /note= "PTK conserved amino acid"
FT   Region          138..145
FT                   /note= "PTK conserved region"
FT   Region          158..160
FT                   /note= "PTK conserved region"
FT   Region          174
FT                   /note= "PTK conserved amino acid"
FT   Region          183..190
FT                   /note= "PTK conserved region"
FT   Region          202..207
FT                   /note= "PTK conserved region"
FT   Region          253
FT                   /note= "PTK conserved amino acid"
FT   Region          256..257
FT                   /note= "PTK conserved region"
FT   Region          264..265
FT                   /note= "PTK conserved region"
FT   Region          274
FT                   /note= "PTK conserved amino acid"
XX
PN   CA2083521-A.
XX
PD   01-OCT-1993.
XX
PF   23-NOV-1992;   92CA-2083521.
XX
PR   31-MAR-1992;   92US-0861390.
XX
PA   (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
PI   Letwin K,  Pawson A,  Reedijk M;
XX
DR   WPI; 1993-406300/51.
DR   N-PSDB; AAQ53470.
XX
PT   Expression of phosphorylated exogenous protein - in host cells
PT   transformed with two vectors, one for the protein, the other for
PT   catalytic domain of protein kinase
XX
PS   Disclosure; Fig 1; 55pp; English.
XX
CC   This sequence is encoded by a fragment of the lambda gt11 expression
CC   vector, lambda-B1-Elk, and represents the catalytic sequence of the
CC   protein tyrosine kinase, Elk.  The Elk gene, B1, encodes a protein
CC   which is a member of the Eph subfamily of protein tyrosine kinases.
CC   The Elk product is very similar to two other receptor-like tyrosine
CC   kinases, eph and eck.  Lambda-B1-Elk may be used in the production
CC   of phosphorylated exogenous protein along with a further vector
CC   encoding the desired exogenous protein.  These plasmid may be used
CC   to produce phosphorylated proteins in host cells which have no
CC   intrinsic capacity for phosphorylation, eg. bacteria.  The system
CC   may be used for the expression of the phosphorylated kinase insert
CC   domain of a growth factor receptor kinase eg. platelet-derived growth
CC   factor receptor.
XX
SQ   Sequence   380 AA;


  Query Match             1.5%;  Score 8;  DB 14;  Length 380;
  Best Local Similarity   100.0%;  Pred. No. 19;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     418 MTSEDLLR 425
           ||||||||
Db     342 mtsedllr 349


RESULT  20
AAW09825
ID   AAW09825 standard; Protein; 466 AA.
XX
AC   AAW09825;
XX
DT   15-JUL-1997  (first entry)
XX
DE   UDP-glucose:thiohydroximate S-glucosyltransferase.
XX
KW   Glucosinolate; UDP-glucose:thiohydroximate S-glucosyltransferase;
KW   S-GT; transgenic plant; rapeseed oil; oilseed rape; canola.
XX
OS   Brassica napus cv. Westar.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "residue 2 is Val in other S-GT isoforms"
FT   Misc-difference 10..11
FT                   /note= "a Lys residue is inserted between amino
```

```
FT                      acids 10 and 11 in some S-GT isoforms"
FT   Misc-difference 12
FT                   /note= "residue 12 is Ser in some S-GT isoforms"
FT   Misc-difference 43
FT                   /note= "residue 43 is Leu in some S-GT isoforms"
FT   Misc-difference 75
FT                   /note= "residue 75 is Pro in some S-GT isoforms"
FT   Misc-difference 88
FT                   /note= "residue 88 is Gly in some S-GT isoforms"
FT   Misc-difference 93
FT                   /note= "residue 93 is His in some S-GT isoforms"
FT   Misc-difference 96
FT                   /note= "residue 96 is Gln in some S-GT isoforms"
FT   Misc-difference 133
FT                   /note= "residue 133 is Leu in some S-GT isoforms"
FT   Misc-difference 153
FT                   /note= "residue 153 is Ala in some S-GT isoforms"
FT   Misc-difference 167
FT                   /note= "residue 167 is Leu in some S-GT isoforms"
FT   Misc-difference 204
FT                   /note= "residue 204 is Ile in some S-GT isoforms"
FT   Misc-difference 216
FT                   /note= "residue 216 is Gly in some S-GT isoforms"
FT   Misc-difference 232
FT                   /note= "residue 232 is Thr in some S-GT isoforms"
FT   Misc-difference 234
FT                   /note= "residue 234 is Lys in some S-GT isoforms"
FT   Misc-difference 243
FT                   /note= "residue 243 is Asp in some S-GT isoforms"
FT   Misc-difference 249
FT                   /note= "residue 249 is Ala in some S-GT isoforms"
FT   Misc-difference 290
FT                   /note= "residue 290 is Arg in some S-GT isoforms"
FT   Misc-difference 302
FT                   /note= "residue Thr is Leu in some S-GT isoforms"
FT   Misc-difference 319
FT                   /note= "residue 319 is Arg in some S-GT isoforms"
FT   Misc-difference 350
FT                   /note= "residue 350 is Glu or Gly in some S-GT
FT                    isoforms"
FT   Misc-difference 395
FT                   /note= "residue 395 is Lys in some S-GT isoforms"
FT   Misc-difference 402
FT                   /note= "residue 402 is Asp in some S-GT isoforms"
FT   Misc-difference 419
FT                   /note= "residue 419 is Lys in some S-GT isoforms"
XX
PN   EP771878-A1.
XX
PD   07-MAY-1997.
XX
PF   31-OCT-1995;   95EP-0402425.
XX
PR   31-OCT-1995;   95EP-0402425.
XX
PA   (CANA ) NAT RES COUNCIL CANADA.
PA   (PLBZ ) PLANT GENETIC SYSTEMS NV.
XX
PI   Grootwassink JWD,  Hemmingsen SM,  Kolenovsky AD,  Peferoen M;
PI   Reed DW, Underhill EW, Van Audenhove K;
XX
DR   WPI; 1997-247418/23.
DR   N-PSDB; AAT66166.
XX
PT   Plants genetically transformed to interfere with
PT   UDP-glucose:thiohydroximate S-glucosyltransferase gene expression
PT   - useful for production or rapeseed oil with reduced glucosinolate
PT   content
XX
PS   Claim 9; Page 23-25; 35pp; English.
XX
CC   A UDP-glucose:thiohydroximate S-glucosyltransferase (S-GT) (AAW09825)
CC   is encoded by clone pGL9 (AAT66166) amplified from Brassica napus cv.
CC   Westar cDNA.  S-GT is the enzyme responsible for the biosynthesis
CC   of glucosinolate.  Novel chimeric genes encode an antisense RNA
CC   complementary to all or part of an mRNA, the cDNA of which is
CC   contained in pGL9.  Oilseed rape plants transformed with these
CC   chimeric genes have reduced contents of glucosinolates, pref.
CC   alkenyl glucosinolates.  This allows the prodn. of rapeseed oil
CC   with a low glucosinolate content.
XX
SQ   Sequence   466 AA;


  Query Match             1.5%;  Score 8;  DB 18;  Length 466;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     385 GVPMVGVP 392
           ||||||||
Db     371 gvpmvgvp 378


RESULT  21
AAR75704
ID   AAR75704 standard; Protein; 951 AA.
XX
AC   AAR75704;
XX
DT   11-NOV-1995  (first entry)
XX
DE   Eph-related CEK6.
XX
KW   Cek6; Eph; protein tyrosine-kinase; PTK; cancer; diagnosis;
KW   prognosis.
XX
OS   Gallus sp.
XX
FH   Key             Location/Qualifiers
FT   Domain          426..444
FT                   /label= Extracellular_domain
XX
PN   WO9515375-A.
XX
PD   08-JUN-1995.
XX
PF   07-SEP-1994;   94WO-US10140.
XX
PR   03-DEC-1993;   93US-0162809.
XX
PA   (LJOL-) LA JOLLA CANCER RES FOUND.
XX
PI   Pasquale EB,  Sajjadi FG;
XX
DR   WPI; 1995-215256/28.
DR   N-PSDB; AAQ90652.
XX
PT   Eph-related protein tyrosine kinase(s) - for monitoring and diagnosing
PT   cancer.
XX
PS   Claim 12; Page 37-41; 129pp; English.
XX
CC   Novel EPH-related PTK Cek6 cDNA clones (AAQ90652) were isolated from
CC   chick embryo and embryonic brain cDNA libraries in phage lambda gt11.
CC   The encoded Cek6 protein (AAR75704) is closely related to rat Elk,
CC   Cek5 (AAR75712) and Cek10 (AAR75708).  Cek6 transcripts were found in
CC   10-day embryos and in adult brain, lung, heart and skeletal muscle.
XX
SQ   Sequence   951 AA;


  Query Match             1.5%;  Score 8;  DB 16;  Length 951;
  Best Local Similarity   100.0%;  Pred. No. 45;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     418 MTSEDLLR 425
```

```
          ||||||||
Db    913 mtsedllr 920


RESULT  22
AAR44513
ID   AAR44513 standard; Protein; 984 AA.
XX
AC   AAR44513;
XX
DT   16-JUN-1994  (first entry)
XX
DE   elk.
XX
KW   Lambda gt11; expression vector; lambda-B1-Elk; protein tyrosine kinase;
KW   Elk; B1; Eph; subfamily; receptor-like tyrosine kinase; eph; eck;
KW   phosphorylation; phosphorylated kinase insert domain; growth factor;
KW   receptor kinase; platelet-derived growth factor receptor.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..17
FT                   /note= "Signal peptide"
FT   Misc-difference 61
FT                   /note= "Cysteine residue"
FT   Misc-difference 96
FT                   /note= "Cysteine residue"
FT   Misc-difference 106
FT                   /note= "Cysteine residue"
FT   Misc-difference 183
FT                   /note= "Cysteine residue"
FT   Misc-difference 196
FT                   /note= "Cysteine residue"
FT   Misc-difference 225
FT                   /note= "Cysteine residue"
FT   Misc-difference 240
FT                   /note= "Cysteine residue"
FT   Misc-difference 253
FT                   /note= "Cysteine residue"
FT   Misc-difference 255
FT                   /note= "Cysteine residue"
FT   Misc-difference 267
FT                   /note= "Cysteine residue"
FT   Misc-difference 270
FT                   /note= "Cysteine residue"
FT   Misc-difference 284
FT                   /note= "Cysteine residue"
FT   Misc-difference 287
FT                   /note= "Cysteine residue"
FT   Misc-difference 301
FT                   /note= "Cysteine residue"
FT   Misc-difference 303
FT                   /note= "Cysteine residue"
FT   Misc-difference 319
FT                   /note= "Cysteine residue"
FT   Misc-difference 360
FT                   /note= "Cysteine residue"
FT   Misc-difference 363
FT                   /note= "Cysteine residue"
FT   Misc-difference 370
FT                   /note= "Cysteine residue"
FT   Misc-difference 373
FT                   /note= "Cysteine residue"
FT   Modified-site   425..427
FT                   /note= "N-glycosylation site"
FT   Modified-site   480..482
FT                   /note= "N-glycosylation site"
XX
PN   CA2083521-A.
XX
PD   01-OCT-1993.
XX
PF   23-NOV-1992;   92CA-2083521.
XX
PR   31-MAR-1992;   92US-0861390.
XX
PA   (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
PI   Letwin K,  Pawson A,  Reedijk M;
XX
DR   WPI; 1993-406300/51.
DR   N-PSDB; Q753471.
XX
PT   Expression of phosphorylated exogenous protein - in host cells
PT   transformed with two vectors, one for the protein, the other for
PT   catalytic domain of protein kinase
XX
PS   Disclosure; Fig 3; 55pp; English.
XX
CC   This sequence is encoded by the elk cNDA and represents the protein
CC   tyrosine kinase, Elk.  The Elk gene, B1, encode a protein which is
CC   a member of the Eph subfamily of protein tyrosine kinases.  The Elk
CC   product is very similar to two other receptor-like tyrosine kinases,
CC   eph and eck.  Lambda-B1-Elk may be used in the production of
CC   phosphorylated exogenous protein along with a further vector encoding
CC   the desired exogenous protein.  These plasmid may be used to produce
CC   phosphorylated proteins in host cells which have no intrinsic capacity
CC   for phosphorylation, eg. bacteria.  The system may be used for the
CC   expression of the phosphorylated kinase insert domain of a growth
CC   factor receptor kinase eg. platelet-derived growth factor receptor.
XX
SQ   Sequence   984 AA;


  Query Match             1.5%;  Score 8;  DB 14;  Length 984;
  Best Local Similarity   100.0%;  Pred. No. 46;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     418 MTSEDLLR 425
           ||||||||
Db     946 mtsedllr 953


RESULT  23
AAR26152
ID   AAR26152 standard; Protein; 17 AA.
XX
AC   AAR26152;
XX
DT   27-JAN-1993  (first entry)
XX
DE   Transferase conserved motif.
XX
KW   Bilirubin; UDP-glucuronic acid.
XX
OS   Synthetic.
XX
PN   WO9212987-A.
XX
PD   06-AUG-1992.
XX
PF   10-JAN-1992;   92WO-US00282.
XX
PR   10-JAN-1991;   91US-0639453.
XX
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI   Owens IS,  Ritter JK;
XX
DR   WPI; 1992-284593/34.
DR   N-PSDB; AAQ27365.
XX
PT   Isolated gene locus UGT1, DNA segments and diagnostic probes -
PT   for diagnosing Gilbert's disease and Crigler-Najjar syndrome
PT   types I and II
```

```
XX
PS   Example 5; Page 40; 99pp; English.
XX
CC   In order to design a universal probe to detect all transferase cDNA
CC   including that which is specific for bilirubin, all characterised
CC   transferase clones were analysed for conserved sequences.  It was
CC   found that a 56 bp sequence is present near the carboxy terminus
CC   of all transferase examined.  The probe of AAQ27365 corresp. to a 90%
CC   conserved motif comprising the sequence below was used.  This 17
CC   amino acid sequence starting with His at position 481 is located
CC   in the lumen of the endoplasmic reticulum and overlaps the membrane-
CC   spanning region by 6 residues.  It is speculated that this sequence,
CC   preceded by a conserved Arg at -4 residues, is a possible binding
CC   site for the common donor substrate, UDP-glucuronic acid.
XX
SQ   Sequence   17 AA;


  Query Match             1.3%;  Score 7;  DB 13;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    479 HDLTWFQ 485
          |||||||
Db      1 hdltwfq 7


RESULT  24
AAR25213
ID   AAR25213 standard; Protein; 30 AA.
XX
AC   AAR25213;
XX
DT   23-DEC-1992  (first entry)
XX
DE   Immunosuppressive peptide anologue #3 of apoE.
XX
KW   inhibit lymphocyte proliferation; ovarian androgen secretion;
KW   ovaries; low density lipoprotein receptor; LDL; steroidogenesis;
KW   hepatic LDL-binding; autoimmune diseases; arthritis;
KW   polycystic ovaries; hypercholesterolaemia.
XX
OS   Synthetic.
XX
PN   WO9210512-A.
XX
PD   25-JUN-1992.
XX
PF   10-DEC-1991;   91WO-US09269.
XX
PR   10-DEC-1990;   90US-0625093.
PR   30-SEP-1991;   91US-0769629.
PR   09-DEC-1991;   91US-0805193.
XX
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Curtiss LK,  Dyer CA,  Smith R;
XX
DR   WPI; 1992-234586/28.
XX
PT   Immunosuppressive polypeptide analogues of apolipoprotein E - for
PT   modulating lymphocyte proliferation and ovarian androgen
PT   synthesis, e.g. for treating inflammation, polycystic ovaries,
PT   hypercholesterolaemia, and in diagnosis
XX
PS   Claim 4; Page 105; 118pp; English.
XX
CC   This peptide is made up of segments corresponding to residues
CC   141-150 of mature apoE which defines a site on apoE involved in low
CC   density lipoprotein (LDL)-receptor binding. The peptides may be used
CC   to modulate physiological events induced by native apoE such as
CC   immune response, steroidogenesis, and/or enhance hepatic LDL-binding.
CC   They may also be used to inhibit lymphocyte proliferation eg in
CC   autoimmune diseases such as in arthritic inflammation. They can also
CC   be used to inhibit ovarian androgen production such as in females
CC   having polycystic ovaries. At lower concentrations the peptides can
CC   enhance lymphocyte proliferation and androgen synthesis. They may
CC   also be used to modulate hepatic LDL binding and uptake, as in
CC   hypercholesterolaemia. Antibodies to the peptides can be used to
CC   monitor the fate of therapeutically administered apoE peptides or
CC   to detect levels of apoE in body samples. See also AAR25211-26.
XX
SQ   Sequence   30 AA;


  Query Match             1.3%;  Score 7;  DB 13;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    422 DLLRALR 428
          |||||||
Db     14 dllralr 20


RESULT  25
AAY07220
ID   AAY07220 standard; Protein; 81 AA.
XX
AC   AAY07220;
XX
DT   16-JUL-1999  (first entry)
XX
DE   Presenilin/beta-amyloid protein interaction inhibitor protein #2.
XX
KW   Inhibition; interaction; beta-amyloid peptide; neurodegeneration;
KW   presenilin; Alzheimer's disease; ligand.
XX
OS   Homo sapiens.
XX
PN   WO9921886-A1.
XX
PD   06-MAY-1999.
XX
PF   23-OCT-1998;   98WO-FR02278.
XX
PR   07-AUG-1998;   98US-0095671.
PR   24-OCT-1997;   97FR-0013384.
XX
PA   (RHON ) RHONE-POULENC RORER SA.
XX
PI   Czech C,  Mercken L,  Pradier L,  Reboul-Becquart S;
XX
DR   WPI; 1999-302983/25.
DR   N-PSDB; AAX57706.
XX
PT   Polypeptides that inhibit interaction of presenilin with amyloid
PT   peptide useful for treating neurodegeneration
XX
PS   Claim 6; Page 84; 101pp; French.
XX
CC   This sequence represents a protein able to inhibit the interaction
CC   between a presenilin (PS) and a beta-amyloid peptide (beta-A) and/or
CC   its precursor, amyloid precursor protein (APP).  The nucleic acid and
CC   protein sequences are used to treat neurodegeneration, particularly
CC   Alzheimer's disease. The proteins are also used to detect cognate
CC   ligands, ligands for PS, beta-A or APP, and compounds that inhibit
CC   interaction of PS with beta-A or APP.
XX
SQ   Sequence   81 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 81;
  Best Local Similarity   100.0%;  Pred. No. 44;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    185 LPAPLSY 191
```

```
          |||||||
Db      4 lpaplsy 10


RESULT  26
AAG01230
ID   AAG01230 standard; Protein; 86 AA.
XX
AC   AAG01230;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein, SEQ ID NO: 5311.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC01236.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 13; SEQ ID 5311; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
CC   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC   were prepared from total human RNAs or polyA+ RNAs derived from 30
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   86 AA;


  Query Match             1.3%;  Score 7;  DB 21;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 47;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    185 LPAPLSY 191
          |||||||
Db      4 lpaplsy 10


RESULT  27
AAB76802
ID   AAB76802 standard; Protein; 88 AA.
XX
AC   AAB76802;
XX
DT   11-APR-2001  (first entry)
XX
DE   Corynebacterium glutamicum MCT protein SEQ ID NO:586.
XX
KW   Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
KW   membrane construction and membrane transport protein; petroleum spill;
KW   hydrocarbon degradation; gram positive aerobic bacterium; marker;
KW   identification; microorganism; fine chemical production; transformation;
KW   genome mapping; genetic engineering.
XX
OS   Corynebacterium glutamicum.
XX
PN   WO200100805-A2.
XX
PD   04-JAN-2001.
XX
PF   23-JUN-2000; 2000WO-IB00926.
XX
PR   25-JUN-1999;   99US-0141031.
PR   08-JUL-1999;   99DE-1031454.
PR   08-JUL-1999;   99DE-1031478.
PR   08-JUL-1999;   99DE-1031563.
PR   09-JUL-1999;   99DE-1032122.
PR   09-JUL-1999;   99DE-1032124.
PR   09-JUL-1999;   99DE-1032125.
PR   09-JUL-1999;   99DE-1032128.
PR   09-JUL-1999;   99DE-1032180.
PR   09-JUL-1999;   99DE-1032182.
PR   09-JUL-1999;   99DE-1032190.
PR   09-JUL-1999;   99DE-1032191.
PR   09-JUL-1999;   99DE-1032209.
PR   09-JUL-1999;   99DE-1032212.
PR   09-JUL-1999;   99DE-1032227.
PR   09-JUL-1999;   99DE-1032228.
PR   09-JUL-1999;   99DE-1032229.
PR   09-JUL-1999;   99DE-1032230.
PR   14-JUL-1999;   99DE-1032927.
PR   14-JUL-1999;   99DE-1033005.
PR   14-JUL-1999;   99DE-1033006.
PR   27-AUG-1999;   99DE-1040764.
PR   27-AUG-1999;   99DE-1040765.
PR   27-AUG-1999;   99DE-1040766.
PR   27-AUG-1999;   99DE-1040830.
PR   27-AUG-1999;   99DE-1040831.
PR   27-AUG-1999;   99DE-1040832.
PR   27-AUG-1999;   99DE-1040833.
PR   31-AUG-1999;   99DE-1041378.
PR   31-AUG-1999;   99DE-1041379.
PR   31-AUG-1999;   99DE-1041395.
PR   03-SEP-1999;   99DE-1042077.
PR   03-SEP-1999;   99DE-1042078.
PR   03-SEP-1999;   99DE-1042079.
PR   03-SEP-1999;   99DE-1042088.
XX
PA   (BADI ) BASF AG.
XX
PI   Pompejus M,  Kroeger B,  Schroeder H,  Zelder O,  Haberhauer G;
XX
DR   WPI; 2001-071486/08.
DR   N-PSDB; AAF68035.
XX
PT   Corynebacterium glutamicum nucleic acids encoding membrane construction
PT   and membrane transport proteins or their portions, useful for typing or
PT   identifying C. glutamicum or related bacteria, and as markers for
PT   transformation -
XX
PS   Claim 20; Page 985; 1119pp; English.
XX
CC   AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC   construction and membrane transport (MCT) proteins given in AAB76510 to
CC   AAB76847. The MCT nucleic acids and proteins are useful in the
CC   identification of microorganisms which can be used to produce fine
CC   chemicals, for modulating fine chemical production in C. glutamicum or
CC   related bacteria (e.g. Brevibacterium lactofermentum), the typing or
```

CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention.
XX
SQ Sequence 88 AA;

Query Match 1.3%; Score 7; DB 22; Length 88;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 ALVFLLL 13
||||||| |
Db 30 alvflll 36

RESULT 28
AAB38042
ID AAB38042 standard; Peptide; 89 AA.
XX
AC AAB38042;
XX
DT 31-JAN-2001 (first entry)
XX
DE Fragment of human secreted protein encoded by gene 10 clone HWHGP71.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200055371-A1.
XX
PD 21-SEP-2000.
XX
PF 16-MAR-2000; 2000WO-US06783.
XX
PR 18-MAR-1999; 99US-0125055.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ni J, Ebner R, Rosen CA, Shi Y, Birse C, Florence K;
PI Komatsoulis G, Lafleur DW, Moore PA, Olsen HS, Young PE;
XX
DR WPI; 2000-594448/56.
XX
PT New nucleic acid molecules encoding 27 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Disclosure; Page 27; 453pp; English.
XX
CC Sequences AAB37984-B38019 represent the amino acid sequences of 27
CC human secreted proteins encoded by the genes AAC69084-C69119. The genes
CC and proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX
SQ Sequence 89 AA;

Query Match 1.3%; Score 7; DB 21; Length 89;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 SKALGRP 241
||||||| |
Db 78 skalgrp 84

RESULT 29
AAB11484
ID AAB11484 standard; peptide; 99 AA.
XX
AC AAB11484;
XX
DT 02-MAR-2001 (first entry)
XX
DE Human presenilin peptide fragment.
XX
KW Protease; membrane-associated substrate; gamma-secretase; presenilinase;
KW NOTCH protease; presenilin-cleaving caspase; protease inhibitor;
KW neurodegenerative disease; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN DE19920514-A1.
XX
PD 16-NOV-2000.
XX
PF 05-MAY-1999; 99DE-1020514.
XX
PR 05-MAY-1999; 99DE-1020514.
XX
PA (BOEH) BOEHRINGER INGELHEIM PHARMA KG.
XX
PI Haass C, Steiner H, Pesold B;
XX
DR WPI; 2001-000492/01.
XX
PT Identifying protease that cleaves membrane-associated substrate, useful
PT e.g. for developing specific therapeutic inhibitors -
XX
PS Claim 32; Fig 6; 10pp; German.
XX
CC This invention describes novel proteases (I) that cleave
CC membrane-associated substrates which are identified by expressing, in a
CC suitable system that is associated with a membrane, a recombinant fusion
CC protein (FP) comprising a reporter component (II) and a protease cleavage
CC site (II) and an additional protein (III), then identifying any (II)
CC cleaved form FP. The products of the invention are used to identify (I)
CC and the genes that encode them from gene libraries, especially
CC gamma-secretase, presenilinase, NOTCH protease and presenilin-cleaving
CC caspase. (I) are important in understanding disease at the molecular
CC level and for development of highly specific protease inhibitors for
CC therapeutic use, e.g. against neurodegenerative diseases such as
CC Alzheimer's disease. The method can detect proteases that are active
CC against membrane-bound substrates, known methods are limited to those
CC that act on cytosolic substrates.
XX
SQ Sequence 99 AA;

Query Match 1.3%; Score 7; DB 22; Length 99;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 185 LPAPLSY 191
||||||| |
Db 2 lpaplsy 8

```
RESULT  30
AAY60184
ID   AAY60184 standard; Protein; 140 AA.
XX
AC   AAY60184;
XX
DT   31-JAN-2000  (first entry)
XX
DE   Human endometrium tumour EST encoded protein 244.
XX
KW   Endometrium; human; tumour; cancer; anticancer; cytostatic; EST:
KW   treatment; uterine; gene therapy; expressed sequence tag.
XX
OS   Homo sapiens.
XX
PN   DE19817948-A1.
XX
PD   21-OCT-1999.
XX
PF   17-APR-1998;   98DE-1017948.
XX
PR   17-APR-1998;   98DE-1017948.
XX
PA   (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI   Rosenthal A,  Specht T,  Hinzmann B,  Schmitt A,  Pilarsky C,  Dahl E;
XX
DR   WPI; 1999-591957/51.
DR   N-PSDB; AAZ42061.
XX
PT   New nucleic acid sequences expressed in uterine cancer tissues, and
PT   derived polypeptides, for treatment of uterine and endometrial cancer
PT   and identification of therapeutic agents  -
XX
PS   Claim 23; Page 373; 444pp; German.
XX
CC   This invention describes novel human nucleic acid (cDNA) sequences (A),
CC   that are highly expressed in uterine tumour tissue and which have
CC   anticancer and cytostatic activity. (A) are used (i) for recombinant
CC   expression of polypeptides (B) and (ii) to isolate complete genes. (B)
CC   are used (i) to identify agents suitable for treatment of uterine or
CC   endometrial cancer; (ii) directly for treating these forms of cancer
CC   (including expression from gene therapy vectors) and (iii) for
CC   generation of specific antibodies. (A) are identified by assembling ESTs
CC   (expressed sequence tags) from a particular tissue type before comparison
CC   of expression patterns. This allows a significantly longer fragment of
CC   the gene to be revealed, so should reduce the number of failures
CC   associated with the fact that ESTs from different libraries may represent
CC   different parts of the same unknown gene, distorting the estimated
CC   frequency of occurrence in a particular tissue. AAY59941-Y60328 represent
CC   protein fragments encoded by the human endometrium tumour cDNA library
CC   derived EST fragments represented in AAZ41981-Z42121.
XX
SQ   Sequence   140 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 140;
  Best Local Similarity   100.0%;  Pred. No. 74;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     151 VIPCGDL 157
           |||||||
Db      16 vipcgdl 22


RESULT  31
AAB32877
ID   AAB32877 standard; Protein; 176 AA.
XX
AC   AAB32877;
XX
DT   25-JAN-2001  (first entry)
XX
DE   Pinus radiata transcription factor protein sequence #4.
XX
KW   Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW   poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW   basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW   homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW   type 2 Cys2His2; CCAAT box element; MYB.
XX
OS   Pinus radiata.
XX
PN   WO200053724-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US06112.
XX
PR   11-MAR-1999;   99US-0266513.
PR   18-AUG-1999;   99US-0149485.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI   Wood M,  McGrath A,  Shenk MA,  Glenn M;
XX
DR   WPI; 2000-579369/54.
XX
PT   New isolated polynucleotide encoding a plant transcription factor for
PT   producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT   having modified gene expression or modified activity of a polypeptide
PT   -
XX
PS   Claim 8; Page 338; 747pp; English.
XX
CC   The present invention relates to novel plant transcription factors from
CC   Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC   transcription factor. The transcription factor may be used to produce a
CC   plant having modified gene expression such as a woody plant e.g. a
CC   eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC   to modify the activity of a polypeptide in a plant. The transcription
CC   factors of the present invention are members from the following families
CC   of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC   helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC   zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
CC   Cys2His2, CCAAT box elements and MYB.
XX
SQ   Sequence   176 AA;


  Query Match             1.3%;  Score 7;  DB 21;  Length 176;
  Best Local Similarity   100.0%;  Pred. No. 91;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     405 KAKGAAV 411
           |||||||
Db     121 kakgaav 127


RESULT  32
AAB33265
ID   AAB33265 standard; Protein; 176 AA.
XX
AC   AAB33265;
XX
DT   25-JAN-2001  (first entry)
XX
DE   Pinus radiata transcription factor protein sequence #311.
XX
KW   Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW   poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW   basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW   homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW   type 2 Cys2His2; CCAAT box element; MYB.
```

```
XX
OS   Pinus radiata.
XX
PN   WO200053724-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US06112.
XX
PR   11-MAR-1999;   99US-0266513.
PR   18-AUG-1999;   99US-0149485.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI   Wood M,  McGrath A,  Shenk MA,  Glenn M;
XX
DR   WPI; 2000-579369/54.
XX
PT   New isolated polynucleotide encoding a plant transcription factor for
PT   producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT   having modified gene expression or modified activity of a polypeptide
PT   -
XX
PS   Claim 8; Page 700; 747pp; English.
XX
CC   The present invention relates to novel plant transcription factors from
CC   Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC   transcription factor. The transcription factor may be used to produce a
CC   plant having modified gene expression such as a woody plant e.g. a
CC   eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC   to modify the activity of a polypeptide in a plant. The transcription
CC   factors of the present invention are members from the following families
CC   of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC   helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC   zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
CC   Cys2His2, CCAAT box elements and MYB.
XX
SQ   Sequence   176 AA;


  Query Match             1.3%;  Score 7;  DB 21;  Length 176;
  Best Local Similarity   100.0%;  Pred. No. 91;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     405 KAKGAAV 411
           |||||||
Db     121 kakgaav 127


RESULT  33
AAY68872
ID   AAY68872 standard; Protein; 180 AA.
XX
AC   AAY68872;
XX
DT   16-MAY-2000  (first entry)
XX
DE   Amino acid sequence of a human presenilin-associated protein HPAP-1.
XX
KW   Human; presenilin-associated protein; HPAP-1;. Incyte clone 1353337;
KW   neurological disorder; cancer; immune disorder; reproductive disorder;
KW   gene therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   70
FT                   /note= "potential phosphorylation site for protein
FT                         kinase C"
FT   Modified-site   95
FT                   /note= "potential phosphorylation site for protein
FT                         kinase C"
FT   Domain          85..101
FT                   /note= "potential transmembrane domain"
FT   Modified-site   103
FT                   /note= "potential phosphorylation site for casein
FT                         kinase II or protein kinase C"
FT   Domain          130..148
FT                   /note= "potential transmembrane domain"
FT   Modified-site   158
FT                   /note= "potential phosphorylation site for protein
FT                         kinase C"
XX
PN   WO200004150-A1.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15858.
XX
PR   16-JUL-1998;   98US-0116640.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Corley NC,  Patterson C;
XX
DR   WPI; 2000-182420/16.
DR   N-PSDB; AAZ46199.
XX
PT   Novel human presenilin-associated protein and polynucleotide used in
PT   the diagnosis, treatment and prevention of cancer, and immune,
PT   neurological and reproductive disorders  -
XX
PS   Claim 1; Fig 1A-C; 68pp; English.
XX
CC   The present sequence represents a human presenilin-associated protein,
CC   designated HPAP-1. The HPAP-1 nucleic acids were first identified in
CC   Incyte clone 1353337 from the heart atrium myxoma cDNA library. HPAP-1
CC   has structural and chemical similarity with human presenilin-1-463.
CC   HPAP-1 polynucleotides, polypeptides, agonists, antagonists and
CC   antibodies are used for in the diagnosis, treatment and prevention of
CC   neurological disorders, cancers, immune disorders and reproductive
CC   disorders. The HPAP-1 polynucleotide is a source of probes and
CC   primers which bind may be used to detect the polynucleotide in a
CC   sample from a patient. The HPAP-1 polynucleotide may also be
CC   administered as part of a gene therapy regime.
XX
SQ   Sequence   180 AA;


  Query Match             1.3%;  Score 7;  DB 21;  Length 180;
  Best Local Similarity   100.0%;  Pred. No. 93;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     185 LPAPLSY 191
           |||||||
Db       4 lpaplsy 10


RESULT  34
AAB37993
ID   AAB37993 standard; Protein; 211 AA.
XX
AC   AAB37993;
XX
DT   31-JAN-2001  (first entry)
XX
DE   Human secreted protein encoded by gene 10 clone HWHGP71.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein.
XX
OS   Homo sapiens.
```

XX
PN WO200055371-A1.
XX
PD 21-SEP-2000.
XX
PF 16-MAR-2000; 2000WO-US06783.
XX
PR 18-MAR-1999; 99US-0125055.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Ni J, Ebner R, Rosen CA, Shi Y, Birse C, Florence K;
PI Komatsoulis G, Lafleur DW, Moore PA, Olsen HS, Young PE;
XX
DR WPI; 2000-594448/56.
DR N-PSDB; AAC69093.
XX
PT New nucleic acid molecules encoding 27 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
PS Claim 11; Page 405; 453pp; English.
XX
CC Sequences AAB37984-B38019 represent the amino acid sequences of 27
CC human secreted proteins encoded by the genes AAC69084-C69119. The genes
CC and proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
XX
SQ Sequence 211 AA;

Query Match 1.3%; Score 7; DB 21; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 SKALGRP 241
|||||||
Db 203 skalgrp 209

RESULT 35
AAY19973
ID AAY19973 standard; Protein; 235 AA.
XX
AC AAY19973;
XX
DT 19-JUL-1999 (first entry)
XX
DE B. burgdorferi antigenic protein, t861.aa.
XX
KW Antigenic protein; vaccine; Lyme disease; infection; detection.
XX
OS Borrelia burgdorferi.
XX
PN WO9859071-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US12718.
XX
PR 03-SEP-1997; 97US-0057483.
PR 20-JUN-1997; 97US-0050359.
PR 22-JUL-1997; 97US-0053344.
PR 22-JUL-1997; 97US-0053377.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMMUNE INC.
XX
PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX
DR WPI; 1999-189980/16.
DR N-PSDB; AAX61670.
XX
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases
PT caused by Borrelia, particularly Lyme disease
XX
PS Claim 12; Page 143; 275pp; English.
XX
CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.
XX
SQ Sequence 235 AA;

Query Match 1.3%; Score 7; DB 20; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 71 ALKFEVV 77
|||||||
Db 137 alkfevv 143

RESULT 36
AAY36910
ID AAY36910 standard; Protein; 244 AA.
XX
AC AAY36910;
XX
DT 07-OCT-1999 (first entry)
XX
DE Protein involved in intermediate metabolism of sugars and/or cofactors.
XX
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW nongonococcal uretritis; epidymitis; cervicitis; salpingitis;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX
OS Chlamydia trachomatis.
XX
PN WO9928475-A2.
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GEST) GENSET.
XX
PI Griffais R;
XX
DR WPI; 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX

```
PS   Disclosure; Page 772; 1755pp; English.
XX
CC   AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC   of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
CC   vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC   can also be used to control growth of the microorganism. Chlamydia
CC   trachomatis is responsible for a large number of diseases, e.g. eye
CC   diseases such as conventional trachoma, nonendemic trachoma,
CC   paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC   nongonococcal uretritis, epidymitis, cervicitis, salpingitis,
CC   perihepatitis, bartholinitis; pneumopathy in breast feeding infants;
CC   and venereal lymphogranulomatosis. The polypeptides of the invention
CC   may be of use in treating these diseases.
XX
SQ   Sequence   244 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 244;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     303 VVFSLGS 309
           |||||||
Db     153 vvfslgs 159


RESULT  37
AAY19972
ID   AAY19972 standard; Protein; 261 AA.
XX
AC   AAY19972;
XX
DT   19-JUL-1999  (first entry)
XX
DE   B. burgdorferi antigenic protein, f861.aa.
XX
KW   Antigenic protein; vaccine; Lyme disease; infection; detection.
XX
OS   Borrelia burgdorferi.
XX
PN   WO9859071-A1.
XX
PD   30-DEC-1998.
XX
PF   18-JUN-1998;   98WO-US12718.
XX
PR   03-SEP-1997;   97US-0057483.
PR   20-JUN-1997;   97US-0050359.
PR   22-JUL-1997;   97US-0053344.
PR   22-JUL-1997;   97US-0053377.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (MEDI-) MEDIMMUNE INC.
XX
PI   Choi GH,  Erwin AL,  Hanson MS,  Lathigra R;
XX
DR   WPI; 1999-189980/16.
DR   N-PSDB; AAX61669.
XX
PT   New isolated Borrelia burgdorferi nucleic acids  - used to develop
PT   products for the diagnosis, prevention and treatment of diseases
PT   caused by Borrelia, particularly Lyme disease
XX
PS   Claim 12; Page 143; 275pp; English.
XX
CC   This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC   invention, which is suitable for use in a vaccine. The Bb polypeptides
CC   can be used in vaccines for eliciting protective antibodies to members of
CC   the Borrelia genus, particularly for the use against Lyme disease in
CC   humans and animals. They can be used for preventing or attenuating an
CC   infection caused by a member of the Borrelia genus. The products can also
CC   be used for detection of members of the Borrelia genus.
XX
SQ   Sequence   261 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 261;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      71 ALKFEVV 77
           |||||||
Db     163 alkfevv 169


RESULT  38
AAY19903
ID   AAY19903 standard; Protein; 303 AA.
XX
AC   AAY19903;
XX
DT   19-JUL-1999  (first entry)
XX
DE   B. burgdorferi antigenic protein, t210.aa.
XX
KW   Antigenic protein; vaccine; Lyme disease; infection; detection.
XX
OS   Borrelia burgdorferi.
XX
PN   WO9859071-A1.
XX
PD   30-DEC-1998.
XX
PF   18-JUN-1998;   98WO-US12718.
XX
PR   03-SEP-1997;   97US-0057483.
PR   20-JUN-1997;   97US-0050359.
PR   22-JUL-1997;   97US-0053344.
PR   22-JUL-1997;   97US-0053377.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (MEDI-) MEDIMMUNE INC.
XX
PI   Choi GH,  Erwin AL,  Hanson MS,  Lathigra R;
XX
DR   WPI; 1999-189980/16.
DR   N-PSDB; AAX61600.
XX
PT   New isolated Borrelia burgdorferi nucleic acids  - used to develop
PT   products for the diagnosis, prevention and treatment of diseases
PT   caused by Borrelia, particularly Lyme disease
XX
PS   Claim 12; Page 112; 275pp; English.
XX
CC   This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC   invention, which is suitable for use in a vaccine. The Bb polypeptides
CC   can be used in vaccines for eliciting protective antibodies to members of
CC   the Borrelia genus, particularly for the use against Lyme disease in
CC   humans and animals. They can be used for preventing or attenuating an
CC   infection caused by a member of the Borrelia genus. The products can also
CC   be used for detection of members of the Borrelia genus.
XX
SQ   Sequence   303 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 303;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     452 VKPLDRA 458
           |||||||
Db     285 vkpldra 291


RESULT  39
AAY37181
```

```
ID   AAY37181 standard; Protein; 309 AA.
XX
AC   AAY37181;
XX
DT   07-OCT-1999  (first entry)
XX
DE   Protein involved in intermediate metabolism of polypeptides.
XX
KW   Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW   paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW   nongonococcal uretritis; epidymitis; cervicitis; salpingitis;
KW   bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX
OS   Chlamydia trachomatis.
XX
PN   WO9928475-A2.
XX
PD   10-JUN-1999.
XX
PF   27-NOV-1998;   98WO-IB01939.
XX
PR   04-NOV-1998;   98US-0107077.
PR   28-NOV-1997;   97FR-0015041.
PR   17-DEC-1997;   97FR-0016034.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-371125/31.
XX
PT   Genome sequence of Chlamydia trachomatis
XX
PS   Disclosure; Page 957; 1755pp; English.
XX
CC   AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC   of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
CC   vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC   can also be used to control growth of the microorganism. Chlamydia
CC   trachomatis is responsible for a large number of diseases, e.g. eye
CC   diseases such as conventional trachoma, nonendemic trachoma,
CC   paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC   nongonococcal uretritis, epidymitis, cervicitis, salpingitis,
CC   perihepatitis, bartholinitis; pneumopathy in breast feeding infants;
CC   and venereal lymphogranulomatosis. The polypeptides of the invention
CC   may be of use in treating these diseases.
XX
SQ   Sequence   309 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     406 AKGAAVE 412
           |||||||
Db     234 akgaave 240


RESULT  40
AAY19902
ID   AAY19902 standard; Protein; 322 AA.
XX
AC   AAY19902;
XX
DT   19-JUL-1999  (first entry)
XX
DE   B. burgdorferi antigenic protein, f210.aa.
XX
KW   Antigenic protein; vaccine; Lyme disease; infection; detection.
XX
OS   Borrelia burgdorferi.
XX
PN   WO9859071-A1.
XX
PD   30-DEC-1998.
XX
PF   18-JUN-1998;   98WO-US12718.
XX
PR   03-SEP-1997;   97US-0057483.
PR   20-JUN-1997;   97US-0050359.
PR   22-JUL-1997;   97US-0053344.
PR   22-JUL-1997;   97US-0053377.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (MEDI-) MEDIMMUNE INC.
XX
PI   Choi GH,  Erwin AL,  Hanson MS,  Lathigra R;
XX
DR   WPI; 1999-189980/16.
DR   N-PSDB; AAX61599.
XX
PT   New isolated Borrelia burgdorferi nucleic acids  - used to develop
PT   products for the diagnosis, prevention and treatment of diseases
PT   caused by Borrelia, particularly Lyme disease
XX
PS   Claim 12; Page 112; 275pp; English.
XX
CC   This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC   invention, which is suitable for use in a vaccine. The Bb polypeptides
CC   can be used in vaccines for eliciting protective antibodies to members of
CC   the Borrelia genus, particularly for the use against Lyme disease in
CC   humans and animals. They can be used for preventing or attenuating an
CC   infection caused by a member of the Borrelia genus. The products can also
CC   be used for detection of members of the Borrelia genus.
XX
SQ   Sequence   322 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 322;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     452 VKPLDRA 458
           |||||||
Db     304 vkpldra 310


RESULT  41
AAR41346
ID   AAR41346 standard; Protein; 348 AA.
XX
AC   AAR41346;
XX
DT   23-FEB-1994  (first entry)
XX
DE   Human CAR receptor polypeptide.
XX
KW   Constitutive activator of retinoic acid response elements;
KW   therapeutics; treatment; cancer; lung cancer; thyroid disorders;
KW   Graves' disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          11..75
FT                   /note= "DNA binding domain"
FT   Domain          76..348
FT                   /note= "Gene activation domain"
XX
PN   WO9317041-A.
XX
PD   02-SEP-1993.
XX
PF   22-FEB-1993;   93WO-US01559.
XX
```

PR 26-FEB-1992; 92US-0843350.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Baes MI, Moore DD;
XX
DR WPI; 1993-288358/36.
DR N-PSDB; AAQ46131.
XX
PT New constitutive activator of retinoic acid receptor elements - are
PT used for studying expression and for treating e.g. cancers or
PT Graves disease
XX
PS Claim 1; Page 31-32; 43pp; English.
XX
CC The sequence is that of a human CAR (constitutive activator of
CC retinoic acid response elements) receptor polypeptide. The CAR
CC receptor polypeptide can be used to increase retinoic acid receptor
CC expression for treating cancers or for decreasing thyroid hormone
CC receptor function for treating Graves' disease. Antibodies to the
CC polypeptide can also be used for therapy or for monitoring the
CC levels of CAR receptor produced by a mammal.
XX
SQ Sequence 348 AA;

Query Match 1.3%; Score 7; DB 14; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 KGAAVEI 413
|||||||
Db 195 kgaavei 201

RESULT 42
AAW32536
ID AAW32536 standard; Protein; 348 AA.
XX
AC AAW32536;
XX
DT 26-JAN-1998 (first entry)
XX
DE Constitutively active receptor-alpha.
XX
KW Constitutively active receptor alpha; CAR-alpha; receptor;
KW steroid like compound; androstenol; libido;
KW 5-alpha reductase inhibitor.
XX
OS Homo sapiens.
XX
PN WO9636230-A1.
XX
PD 21-NOV-1996.
XX
PF 17-APR-1996; 96WO-US03865.
XX
PR 16-MAY-1995; 95US-0442464.
XX
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
PI Evans RM, Forman BM;
XX
DR WPI; 1997-011750/01.
DR N-PSDB; AAT92305.
XX
PT Modulating activity of isoform of constitutively active receptor -
PT by admin. of steroid-like cpd. such as androstenol
XX
PS Example 2; Page 23-24; 42pp; English.
XX
CC A steroid-like compound has been developed for modulating the activity
CC of an isoform of CAR (constitutively active receptor) or a CAR-like
CC species. The present sequence represents the receptor CAR-alpha. A
CC method, which has been developed for the identification of compounds
CC which modulate the activity of an isoform of CAR or a CAR-like species,
CC involves: (a) contacting host cells containing receptor-encoded DNA and
CC a hormone response element linked to reporter-encoded DNA with a test
CC compound; and (b) determining the effect of the test compound on the
CC level of expression of the reporter. The steroid-like compounds may be
CC used to modulate processes mediated by CAR or a CAR-like species, and
CC are especially useful in increasing the libido (especially in subjects
CC undergoing therapy with a 5alpha-reductase inhibitor).
XX
SQ Sequence 348 AA;

Query Match 1.3%; Score 7; DB 18; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 407 KGAAVEI 413
|||||||
Db 195 kgaavei 201

RESULT 43
AAW93902
ID AAW93902 standard; Protein; 348 AA.
XX
AC AAW93902;
XX
DT 29-JUN-1999 (first entry)
XX
DE Human CAR receptor protein.
XX
KW CAR receptor; constitutive activator of retinoic acid response element;
KW cytostatic; antithyroid; nuclear hormone receptor superfamily; cancer;
KW zinc finger transcription factor; retinoic acid; treatment; lung;
KW Grave's disease; thyroid hormone receptor; human.
XX
OS Homo sapiens.
XX
PN WO9915555-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-US19365.
XX
PR 19-SEP-1997; 97US-0934388.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Baes MI, Choi H, Moore DD;
XX
DR WPI; 1999-254691/21.
DR N-PSDB; AAX23994.
XX
PT CAR (Constitutive Activator of Retinoic Acid Response Elements)
PT receptor polypeptides, used to identify ligands for therapy of
PT Grave's disease and lung cancer
XX
PS Claim 20; Fig 1; 66pp; English.
XX
CC This invention describes the isolation of novel human and mouse CAR
CC (Constitutive Activator of Retinoic Acid Response Elements) receptor
CC polypeptides which have cytostatic and antithyroid activity. The CAR
CC receptor polypeptides are members of the nuclear hormone receptor
CC superfamily (zinc finger transcription factors). The CAR receptor
CC polypeptides bind to their target DNA sequence and activate expression of
CC downstream genes, even in the absence of retinoic acid. CAR receptor
CC ligands are useful for treating Grave's disease by decreasing thyroid
CC hormone receptor function, and for treating cancer (especially lung
CC cancer) by increasing retinoic acid receptor expression.
XX
SQ Sequence 348 AA;

```
  Query Match             1.3%;  Score 7;  DB 20;  Length 348;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     407 KGAAVEI 413
           |||||||
Db     195 kgaavei 201


RESULT  44
AAY17872
ID   AAY17872 standard; Protein; 357 AA.
XX
AC   AAY17872;
XX
DT   18-AUG-1999  (first entry)
XX
DE   Mouse nuclear receptor protein nNR4.
XX
KW   Mouse; nuclear receptor protein; nNR4; identification; differentiation;
KW   cell proliferation; regulation; murine.
XX
OS   Mus musculus.
XX
PN   WO9929722-A1.
XX
PD   17-JUN-1999.
XX
PF   11-DEC-1998;   98WO-US26446.
XX
PR   12-DEC-1997;   97US-0068144.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Chen F;
XX
DR   WPI; 1999-385573/32.
DR   N-PSDB; AAX80215.
XX
PT   Novel DNA encoding murine nuclear receptor
XX
PS   Claim 23; Page 29; 47pp; English.
XX
CC   The present sequence is a mouse nuclear receptor protein designated
CC   nNR4. The nNR4 protein is useful in the identification of downstream
CC   target genes and ligands regulating its activity. The nuclear receptor
CC   is involved in the regulation of in vivo cell proliferation and/or cell
CC   development. The nNR4 polynucleotides, expression vectors and host cells
CC   are useful for the recombinant production of the protein.
XX
SQ   Sequence   357 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 357;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     407 KGAAVEI 413
           |||||||
Db     205 kgaavei 211


RESULT  45
AAW93903
ID   AAW93903 standard; Protein; 358 AA.
XX
AC   AAW93903;
XX
DT   29-JUN-1999  (first entry)
XX
DE   Mouse CAR receptor protein.
XX
KW   CAR receptor; constitutive activator of retinoic acid response element;
KW   cytostatic; antithyroid; nuclear hormone receptor superfamily; cancer;
KW   zinc finger transcription factor; retinoic acid; treatment; lung;
KW   Grave's disease; thyroid hormone receptor; mouse.
XX
OS   Mus sp.
XX
PN   WO9915555-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19365.
XX
PR   19-SEP-1997;   97US-0934388.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Baes MI,  Choi H,  Moore DD;
XX
DR   WPI; 1999-254691/21.
DR   N-PSDB; AAX24003.
XX
PT   CAR (Constitutive Activator of Retinoic Acid Response Elements)
PT   receptor polypeptides, used to identify ligands for therapy of
PT   Grave's disease and lung cancer
XX
PS   Claim 2; Fig 2; 66pp; English.
XX
CC   This invention describes the isolation of novel human and mouse CAR
CC   (Constitutive Activator of Retinoic Acid Response Elements) receptor
CC   polypeptides which have cytostatic and antithyroid activity. The CAR
CC   receptor polypeptides are members of the nuclear hormone receptor
CC   superfamily (zinc finger transcription factors). The CAR receptor
CC   polypeptides bind to their target DNA sequence and activate expression of
CC   downstream genes, even in the absence of retinoic acid. CAR receptor
CC   ligands are useful for treating Grave's disease by decreasing thyroid
CC   hormone receptor function, and for treating cancer (especially lung
CC   cancer) by increasing retinoic acid receptor expression.
XX
SQ   Sequence   358 AA;


  Query Match             1.3%;  Score 7;  DB 20;  Length 358;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     407 KGAAVEI 413
           |||||||
Db     205 kgaavei 211


Search completed: August 13, 2001, 13:43:35
Job time: 480 sec
```